

SEQUENCE LISTING

<110> Kim, Jospeh L
Morgenstern, Kurt A
Caron, Paul R
Lin, Chao
Vertex Pharmaceuticals Inc.

<120> CRYSTALS OF HEPATITIS C VIRUS HELICASE OR FRAGMENTS THEREOF
COMPRISING A HELICASE BINDING POCKET

<130> Sequence listing for VPI/97-101 CIP CON

<140>

<141>

<150> PCT/US98/16879

<151> 1998-08-13

<160> 4

<170> PatentIn Ver. 2.0

<210> 1

<211> 1932

<212> DNA

<213> Hepatitis C virus

<220>

<221> CDS

<222> (4)...(1896)

<223> Full length HCV NS3 coding sequence

<220>

<221> misc_feature

<220>

<221> misc_feature

<222> (504)..(1896)

<223> Helicase domain

<400> 1

atg gcg ccc atc acg gcg tac gcc cag cag acg aga ggc aag ctt ggg 48

Ala Pro Ile Thr Ala Tyr Ala Gln Gln Thr Arg Gly Lys Leu Gly

1 5 10 15

tgt ata atc acc agc ctg act ggc cgg gac aaa aac caa gtg gag ggt 96

Cys Ile Ile Thr Ser Leu Thr Gly Arg Asp Lys Asn Gln Val Glu Gly

20 25 30

gag gtc cag atc gtg tca act gct acc caa acc ttc ctg gca acg tgc 144

Glu Val Gln Ile Val Ser Thr Ala Thr Gln Thr Phe Leu Ala Thr Cys

35 40 45

atc aat ggg gta tgc tgg act gtc tac cac ggg gcc gga acg agg acc 192

Ile Asn Gly Val Cys Trp Thr Val Tyr His Gly Ala Gly Thr Arg Thr

50 55 60

atc gca tca ccc aag ggt cct gtc atc cag atg tat acc aat gtg gac 240

Ile Ala Ser Pro Lys Gly Pro Val Ile Gln Met Tyr Thr Asn Val Asp

65	70	75	
caa gac ctt gtg ggc tgg ccc gct cct caa ggt tcc cgc tca ttg aca 288			
Gln Asp Leu Val Gly Trp Pro Ala Pro Gln Gly Ser Arg Ser Leu Thr			
80	85	90	95
ccc tgc acc tgc ggc tcc tcg gac ctt tac ctg gtt acg agg cac gcc 336			
Pro Cys Thr Cys Gly Ser Ser Asp Leu Tyr Leu Val Thr Arg His Ala			
100	105	110	
gac gtc atc ccg gtt cgc cgt cgc ggt gat agc cgt ggt agc ctg ctg 384			
Asp Val Ile Pro Val Arg Arg Arg Gly Asp Ser Arg Gly Ser Leu Leu			
115	120	125	
tct ccg cgt ccg att tcc tac ctg aaa ggc tcc tcg ggg ggt ccg ctg 432			
Ser Pro Arg Pro Ile Ser Tyr Leu Lys Gly Ser Ser Gly Gly Pro Leu			
130	135	140	
ttg tgc ccc gcg gga cac gcc gtg ggc cta ttc agg gcc gcg gtg tgc 480			
Leu Cys Pro Ala Gly His Ala Val Gly Leu Phe Arg Ala Ala Val Cys			
145	150	155	
acc cgt gga gtg gcc aag gcg gtg gac ttt atc cct gtg gag aac ctg 528			
Thr Arg Gly Val Ala Lys Ala Val Asp Phe Ile Pro Val Glu Asn Leu			

160	165	170	175
gag acc acc atg cgt tcc ccg gtg ttc acg gac aac tcc tct cca cca 576			
Glu Thr Thr Met Arg Ser Pro Val Phe Thr Asp Asn Ser Ser Pro Pro			
180	185	190	
gct gtt ccc cag agc ttc cag gtg gcc cac ctg cat gct ccc acc ggc 624			
Ala Val Pro Gln Ser Phe Gln Val Ala His Leu His Ala Pro Thr Gly			
195	200	205	
agt ggt aag agc acc aag gtc ccg gct gcg tac gca gcc cag ggc tac 672			
Ser Gly Lys Ser Thr Lys Val Pro Ala Ala Tyr Ala Ala Gln Gly Tyr			
210	215	220	
aag gtg ttg gtg ctc aac ccc tct gtt gct gca acg ctg ggc ttt ggt 720			
Lys Val Leu Val Leu Asn Pro Ser Val Ala Ala Thr Leu Gly Phe Gly			
225	230	235	
gct tac atg tcc aag gcc cat ggg gtc gat cct aat atc cgc acc ggt 768			
Ala Tyr Met Ser Lys Ala His Gly Val Asp Pro Asn Ile Arg Thr Gly			
240	245	250	255
gtg cgt aca att acc act ggc agc ccc atc acg tac tcc acc tac ggc 816			

Val Arg Thr Ile Thr Thr Gly Ser Pro Ile Thr Tyr Ser Thr Tyr Gly

260

265

270

aag ttc ctt gcc gac ggc ggg tgc tca ggt ggc gct tat gat atc atc 864

Lys Phe Leu Ala Asp Gly Gly Cys Ser Gly Gly Ala Tyr Asp Ile Ile

275

280

285

att tgt gac gag tgc cac tcc acg gat gcc aca tcc atc ttg ggc atc 912

Ile Cys Asp Glu Cys His Ser Thr Asp Ala Thr Ser Ile Leu Gly Ile

290

295

300

ggc act gtc ctt gac caa gca gag act gcg ggg gcg aga ttg gtt gtg 960

Gly Thr Val Leu Asp Gln Ala Glu Thr Ala Gly Ala Arg Leu Val Val

305

310

315

ctc gcc act gct acc cct ccg ggc tcc gtc acg gta ccg cat cct aac 1008

Leu Ala Thr Ala Thr Pro Pro Gly Ser Val Thr Val Pro His Pro Asn

320

325

330

335

atc gag gag gtt gct ctg tcc acc acc gga gag atc cct ttc tac ggc 1056

Ile Glu Glu Val Ala Leu Ser Thr Thr Gly Glu Ile Pro Phe Tyr Gly

340

345

350

aag gct atc ccc ctc gag gtg atc aag ggc ggc cgt cat ctc atc ttc 1104

Lys Ala Ile Pro Leu Glu Val Ile Lys Gly Gly Arg His Leu Ile Phe

355

360

365

tgt cac tca aag aag aag tgc gac gag ctc gcc gcg aag ctg gtc gca 1152

Cys His Ser Lys Lys Lys Cys Asp Glu Leu Ala Ala Lys Leu Val Ala

370

375

380

ttg ggc atc aat gcc gtg gcc tac tac cgc gga ctt gac gtg tct gtc 1200

Leu Gly Ile Asn Ala Val Ala Tyr Tyr Arg Gly Leu Asp Val Ser Val

385

390

395

atc ccg acc agc ggc gat gtt gtc gtc gtg gcg acc gat gct ctc atg 1248

Ile Pro Thr Ser Gly Asp Val Val Val Val Ala Thr Asp Ala Leu Met

400

405

410

415

act ggc ttt acc ggc gac ttc gac tct gtg ata gac tgc aac acg tgt 1296

Thr Gly Phe Thr Gly Asp Phe Asp Ser Val Ile Asp Cys Asn Thr Cys

420

425

430

gtc act cag aca gtc gat ttc agc ctt gac cct acc ttt acc att gag 1344

Val Thr Gln Thr Val Asp Phe Ser Leu Asp Pro Thr Phe Thr Ile Glu

435

440

445

aca acc acg ctc ccc cag gat gct gtc tcc agg act cag cgc cgt ggt 1392

Thr Thr Thr Leu Pro Gln Asp Ala Val Ser Arg Thr Gln Arg Arg Gly

450

455

460

cgt acc ggc cgt ggg aag cca ggc atc tac aga ttt gtg gca ccg ggg 1440

Arg Thr Gly Arg Gly Lys Pro Gly Ile Tyr Arg Phe Val Ala Pro Gly

465

470

475

gag cgc ccc tcc ggc atg ttc gac tcg tcc gtc ctc tgt gag tgc tat 1488

Glu Arg Pro Ser Gly Met Phe Asp Ser Ser Val Leu Cys Glu Cys Tyr

480

485

490

495

gac gcg ggc tgt gct tgg tat gag ctc acg ccg gcg gag act aca gtt 1536

Asp Ala Gly Cys Ala Trp Tyr Glu Leu Thr Pro Ala Glu Thr Thr Val

500

505

510

cgt ctg cgc gcg tac atg aac acc ccg ggg ctt ccc gtg tgc cag gac 1584

Arg Leu Arg Ala Tyr Met Asn Thr Pro Gly Leu Pro Val Cys Gln Asp

515

520

525

cat ctt gaa ttt tgg gag ggc gtc ttt acg ggc ctc acc cat atc gat 1632

His Leu Glu Phe Trp Glu Gly Val Phe Thr Gly Leu Thr His Ile Asp

530

535

540

gcc cac ttt ctg tcc cag aca aag cag agt ggg gag aac ttt cct tac 1680

Ala His Phe Leu Ser Gln Thr Lys Gln Ser Gly Glu Asn Phe Pro Tyr

545

550

555

ctg gta gcg tac caa gcc acc gtg tgc gct cgt gcg caa gcc cct ccg 1728

Leu Val Ala Tyr Gln Ala Thr Val Cys Ala Arg Ala Gln Ala Pro Pro

560

565

570

575

cca tcg tgg gac cag atg tgg aag tgt ttg atc cgc ctt aaa ccc acc 1776

Pro Ser Trp Asp Gln Met Trp Lys Cys Leu Ile Arg Leu Lys Pro Thr

580

585

590

ctc cat ggg cca aca ccg ctc ctg tac cgt ctg ggc gct gtt cag aat 1824

Leu His Gly Pro Thr Pro Leu Leu Tyr Arg Leu Gly Ala Val Gln Asn

595

600

605

gaa gtc acc ctg acg cac cca atc acc aaa tac atc atg aca tgc atg 1872

Glu Val Thr Leu Thr His Pro Ile Thr Lys Tyr Ile Met Thr Cys Met

610

615

620

tcg gcc gac ctg gag gtc gtc acg ggatctggct cgcacatca tcatcatcac 1926

Ser Ala Asp Leu Glu Val Val Thr

625

630

taatag

1932

<210> 2

<211> 631

<212> PRT

<213> Hepatitis C virus

<400> 2

Ala Pro Ile Thr Ala Tyr Ala Gln Gln Thr Arg Gly Lys Leu Gly Cys

1 5 10 15

Ile Ile Thr Ser Leu Thr Gly Arg Asp Lys Asn Gln Val Glu Gly Glu

20 25 30

Val Gln Ile Val Ser Thr Ala Thr Gln Thr Phe Leu Ala Thr Cys Ile

35 40 45

Asn Gly Val Cys Trp Thr Val Tyr His Gly Ala Gly Thr Arg Thr Ile

50 55 60

Ala Ser Pro Lys Gly Pro Val Ile Gln Met Tyr Thr Asn Val Asp Gln

65 70 75 80

Asp Leu Val Gly Trp Pro Ala Pro Gln Gly Ser Arg Ser Leu Thr Pro

85 90 95

Cys Thr Cys Gly Ser Ser Asp Leu Tyr Leu Val Thr Arg His Ala Asp

100 105 110

Val Ile Pro Val Arg Arg Arg Gly Asp Ser Arg Gly Ser Leu Leu Ser

115	120	125
Pro Arg Pro Ile Ser Tyr Leu Lys Gly Ser Ser Gly Gly Pro Leu Leu		
130	135	140
Cys Pro Ala Gly His Ala Val Gly Leu Phe Arg Ala Ala Val Cys Thr		
145	150	155
Arg Gly Val Ala Lys Ala Val Asp Phe Ile Pro Val Glu Asn Leu Glu		
165	170	175
Thr Thr Met Arg Ser Pro Val Phe Thr Asp Asn Ser Ser Pro Pro Ala		
180	185	190
Val Pro Gln Ser Phe Gln Val Ala His Leu His Ala Pro Thr Gly Ser		
195	200	205
Gly Lys Ser Thr Lys Val Pro Ala Ala Tyr Ala Ala Gln Gly Tyr Lys		
210	215	220
Val Leu Val Leu Asn Pro Ser Val Ala Ala Thr Leu Gly Phe Gly Ala		
225	230	235
Tyr Met Ser Lys Ala His Gly Val Asp Pro Asn Ile Arg Thr Gly Val		
245	250	255
Arg Thr Ile Thr Thr Gly Ser Pro Ile Thr Tyr Ser Thr Tyr Gly Lys		
260	265	270
Phe Leu Ala Asp Gly Gly Cys Ser Gly Gly Ala Tyr Asp Ile Ile Ile		
275	280	285

Cys Asp Glu Cys His Ser Thr Asp Ala Thr Ser Ile Leu Gly Ile Gly
290 295 300

Thr Val Leu Asp Gln Ala Glu Thr Ala Gly Ala Arg Leu Val Val Leu
305 310 315 320

Ala Thr Ala Thr Pro Pro Gly Ser Val Thr Val Pro His Pro Asn Ile
325 330 335

Glu Glu Val Ala Leu Ser Thr Thr Gly Glu Ile Pro Phe Tyr Gly Lys
340 345 350

Ala Ile Pro Leu Glu Val Ile Lys Gly Gly Arg His Leu Ile Phe Cys
355 360 365

His Ser Lys Lys Lys Cys Asp Glu Leu Ala Ala Lys Leu Val Ala Leu
370 375 380

Gly Ile Asn Ala Val Ala Tyr Tyr Arg Gly Leu Asp Val Ser Val Ile
385 390 395 400

Pro Thr Ser Gly Asp Val Val Val Val Ala Thr Asp Ala Leu Met Thr
405 410 415

Gly Phe Thr Gly Asp Phe Asp Ser Val Ile Asp Cys Asn Thr Cys Val
420 425 430

Thr Gln Thr Val Asp Phe Ser Leu Asp Pro Thr Phe Thr Ile Glu Thr
435 440 445

Thr Thr Leu Pro Gln Asp Ala Val Ser Arg Thr Gln Arg Arg Gly Arg

450	455	460
Thr Gly Arg Gly Lys Pro Gly Ile Tyr Arg Phe Val Ala Pro Gly Glu		
465	470	475 480
Arg Pro Ser Gly Met Phe Asp Ser Ser Val Leu Cys Glu Cys Tyr Asp		
485	490	495
Ala Gly Cys Ala Trp Tyr Glu Leu Thr Pro Ala Glu Thr Thr Val Arg		
500	505	510
Leu Arg Ala Tyr Met Asn Thr Pro Gly Leu Pro Val Cys Gln Asp His		
515	520	525
Leu Glu Phe Trp Glu Gly Val Phe Thr Gly Leu Thr His Ile Asp Ala		
530	535	540
His Phe Leu Ser Gln Thr Lys Gln Ser Gly Glu Asn Phe Pro Tyr Leu		
545	550	555 560
Val Ala Tyr Gln Ala Thr Val Cys Ala Arg Ala Gln Ala Pro Pro Pro		
565	570	575
Ser Trp Asp Gln Met Trp Lys Cys Leu Ile Arg Leu Lys Pro Thr Leu		
580	585	590
His Gly Pro Thr Pro Leu Leu Tyr Arg Leu Gly Ala Val Gln Asn Glu		
595	600	605
Val Thr Leu Thr His Pro Ile Thr Lys Tyr Ile Met Thr Cys Met Ser		
610	615	620
Ala Asp Leu Glu Val Val Thr		

625

630

<210> 3

<211> 8157

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Genetically
engineered plasmid containing full-length HCV NS3
coding sequence

<400> 3

ttaatacgac tcactatagg ggaattgtga gcggataaca attccctct agaaataatt 60

ttgtttaact ttaagaagga gatataccat ggcgcccac acggcgtagc cccagcagac 120

gagaggcaag cttgggtgta taatcaccag cctgactggc cgggacaaaa accaagtgga 180

gggtgaggtc cagatcgtgt caactgctac ccaaaccttc ctggcaacgt gcatcaatgg 240

ggtatgctgg actgtctacc acggggccgg aacgaggacc atcgcatcac ccaagggtcc 300

tgtcatccag atgtatacca atgtggacca agaccttggt ggctggcccg ctctcaagg 360

ttcccgtca ttgacaccct gcacctgagg ctctcggac ctttacctgg ttacgaggca 420

cgccgacgtc atcccggttc gccgtcggg tgatagccgt ggtagcctgc tgtctccgcg 480

tccgatttcc tacctgaaag gtcctcggg gggtcgctg ttgtgccccg cgggacacgc 540

cgtgggacct ttcagggccg cgggtgacac ccgtggagtg gccaaaggcg tggactttat 600

ccctgtggag aacctggaga ccacatgag tccccggtg ttacaggaca actcctctcc 660

accagctgtt cccagagct tccaggtggc ccacctgcat gctcccaccg gcagtggtaa 720

gagcaccaag gtcccggtg cgtacgcagc ccagggtac aagggtgttg tgctcaaccc 780

ctctgttgct gcaacgctgg gctttggtgc ttacatgtcc aaggcccatg gggtcgatcc 840

taatatccgc accggtgtgc gtacaattac cactggcagc cccatcacgt actccaccta 900

cggcaagttc cttgccgacg gcgggtgctc aggtggcgct tatgatata tcatttgtga 960

cgagtgccac tccacggatg ccacatccat cttgggcacg ggcactgtcc ttgaccaagc 1020

agagactgcg ggggcgagat tggttgtgct cgccactgct acccctccgg gctccgtcac 1080

ggtaccgcat cctaacatcg aggaggttgc tctgtccacc accggagaga tccctttcta 1140

cggcaaggct atccccctcg aggtgatcaa gggcggccgt catctcatct tctgtcactc 1200

aaagaagaag tgcgacgagc tcgccgcgaa gctggtcgca ttgggcatca atgccgtggc 1260

ctactaccgc ggacttgacg tgtctgtcat cccgaccagc ggcgatgttg tcgtcgtggc 1320

gaccgatgct ctcatgactg gctttaccgg cgacttcgac tctgtgatag actgcaacac 1380

gtgtgtcact cagacagtcg atttcagcct tgaccctacc ttaccattg agacaaccac 1440

gtccccccag gatgctgtct ccaggactca gcgccgtggc cgtaccggcc gtgggaagcc 1500

aggcatctac agatttgtgg caccggggga gcgccctcc ggcattgtcg actcgtccgt 1560

cctctgtgag tgctatgacg cgggctgtgc ttggtatgag ctcacgccgg cggagactac 1620

agttcgtctg cgcgcgtaca tgaacacccc ggggcttccc gtgtgccagg accatcttga 1680

at t t t t g g g a g g g c g t c t t t a c g g g c c t c a c c c a t a t c g a t g c c c a c t t t c t g t c c c a g a c 1740

a a a g c a g a g t g g g g a g a a c t t t c c t t a c c t g g t a g c g t a c c a a g c c a c c g t g t g c g t c g 1800

t g c g c a a g c c c c t c c g c c a t c g t g g g a c c a g a t g t g g a a g t g t t t g a t c c g c c t t a a a c c 1860

c a c c c t c c a t g g g c c a a c a c c g t c c t g t a c c g t c t g g g c g c t g t t c a g a t g a a g t c a c 1920

c c t g a c g c a c c c a a t c a c c a a a t a c a t c a t g a c a t g c a t g t c g g c c g a c c t g g a g g t c g t 1980

c a c g g g a t c t g g c t c g c a t c a t c a t c a t c a t c a c t a a t a g a a t t c g g a t c c g g t g c t a a 2040

c a a a g c c c g a a a g g a a g c t g a g t t g g c t g c t g c c a c c g c t g a g c a a t a a c t a g c a t a a c c 2100

c c t t g g g g c c t c t a a a c g g g t c t t g a g g g g t t t t t t g c t g a a a g g a g g a a c t a t a t c c g g 2160

a t a t c c c g c a a g a g g c c c g g c a g t a c c g g c a t a a c c a a g c c t a t g c c t a c a g c a t c c a g g 2220

g t g a c g g t g c c g a g g a t g a c g a t g a g c g c a t t g t t a g a t t t c a t a c a c g g t g c c t g a c t g 2280

c g t t a g c a a t t t a a c t g t g a t a a a c t a c c g a t t a a a g c t a t c g a t a c c g t c g a c c t c g 2340

agggggggcc cggtagccaa ttgcacctat agtgagtcgt attacgcgcg ctcaactggcc 2400

gtcgtttttac aacgtcgtga ctgggaaaac cctggcggtta cccaacttaa tcgccttgca 2460

gcacatcccc ctttcgccag ctggcgtaat agcgaagagg ccgcaccga tcgcccttcc 2520

caacagttgc gcagcctgaa tggcgaatgg gacgcgcctt gtagcggcgc attaagcgcg 2580

gcgggtgtgg tggttacgcg cagcgtgacc gctacacttg ccagcgcctt agcgccttgc 2640

cctttcgttt tcttcccttc ctttctcgcc acgttcgcgc gctttccccg tcaagctcta 2700

aatcgggggc tcccttttagg gttccgattt agtgctttac ggcacctcga ccccaaaaaa 2760

cttgattagg gtgatgggtc acgtagtggg ccatcgccct gatagacggt ttttcgccct 2820

ttgacgttgg agtccacgtt ctttaatagt ggactcttgt tccaaactgg aacaacactc 2880

aacctatct cggctctattc ttttgattta taagggattt tgccgatttc ggcctattgg 2940

ttaaaaaatg agctgattta acaaaaattt aacgcgaatt ttaacaaaat attaacgctt 3000

acaatttagg tggcactttt cggggaaatg tgcgcggaac ccctatttgt ttatTTTTtct 3060

aaatacattc aaatatgtat ccgctcatga gacaataacc ctgataaatg cttcaataat 3120

attgaaaaag gaagagtatg agtattcaac atttccgtgt cgcccttatt ccctTTTTtg 3180

cggcatTTTTg ccttcctgtt tttgtcacc cagaaacgct ggtgaaagta aaagatgctg 3240

aagatcagtt ggggtgcacga gtgggttaca tcgaactgga tctcaacagc ggtaagatcc 3300

ttgagagttt tcgccccgaa gaacgttttc caatgatgag cacttttaaa gttctgctat 3360

gtggcgcggt attatcccggt attgacgccg ggcaagagca actcggtcgc cgcatacact 3420

attctcagaa tgacttggtt gagtactcac cagtcacaga aaagcatctt acggatggca 3480

tgacagtaag agaattatgc agtgctgcca taaccatgag tgataaacact ggggccaaact 3540

tacttctgac aacgatcgga ggaccgaagg agctaaccgc ttttttgac aacatggggg 3600

atcatgtaac tcgccttgat cgttgggaac cggagctgaa tgaagccata ccaaacgacg 3660

agcgtgacac cacgatgcct gcagcaatgg caacaacggt gcgcaaacta ttaactggcg 3720

aactacttac tctagcttcc cggcaacaat taatagactg gatggaggcg gataaagttg 3780

caggaccact tctgcgctcg gcccttcggt ctggctggtt tattgctgat aaatctggag 3840

ccggtgagcg tgggtctcgc ggtatcattg cagcactggg gccagatggt aagccctccc 3900

gtatcgtagt tatctacacg acggggagtc aggcaactat ggatgaacga aatagacaga 3960

tcgctgagat aggtgcctca ctgattaagc attggtaact gtcagaccaa gtttactcat 4020

atatacttta gattgattta aaacttcatt ttttaatttaa aaggatctag gtgaagatcc 4080

tttttgataa tctcatgacc aaaatccctt aacgtgagtt ttcgttccac tgagcgtcag 4140

accccgtaga aaagatcaaa ggatcttctt gagatccttt ttttctgcgc gtaatctgct 4200

gcttgcaaac aaaaaaacca ccgctaccag cggtggtttg tttgccggat caagagctac 4260

caactctttt tccgaaggta actggcttca gcagagcgca gataccaaat actgtccttc 4320

tagtgtagcc gtagttaggc caccacttca agaactctgt agcaccgcct acatacctcg 4380

ctctgctaata cctgttacca gtggctgctg ccagtggcga taagtcgtgt cttaccgggt 4440

tggactcaag acgatagtta ccgataagg cgcagcggtc gggctgaacg gggggttcgt 4500

gcacacagcc cagcttggag cgaacgacct acaccgaact gagataccta cagcgtgagc 4560

tatgagaaag cgccacgctt cccgaaggga gaaaggcgga caggtatccg gtaagcggca 4620

gggtcggaac aggagagcgc acgagggagc ttccaggggg aaacgcctgg tatctttata 4680

gtcctgtcgg gtttcgccac ctctgacttg agcgtcgatt tttgtgatgc tcgtcagggg 4740

ggcggagcct atggaaaaac gccagcaacg cggccttttt acggttcttg gccttttget 4800

ggccttttgc tcacatgttc tttcctgcgt tatccctga ttctgtggat aaccgtatta 4860

ccgcctttga gtgagctgat accgctcgcc gcagccgaac gaccgagcgc agcgagtcag 4920

tgagcgagga agcgggaagag cgcctgatgc ggtatcttct ccttacgcat ctgtgcggt 4980

tttcacaccg catatatggt gcaactctcag tacaatctgc tctgatgccg catagttaag 5040

ccagtataca ctccgctatc gctacgtgac tgggtcatgg ctgcgccccg acaccgccca 5100

acaccgctg acgcgccctg acgggcttgt ctgctcccg catccgctta cagacaagct 5160

gtgaccgtct ccgggagctg catgtgtcag aggttttcac cgatcatcacc gaaacgcgcg 5220

aggcagctgc ggtaaagctc atcagcgtgg tcgtgaagcg attcacagat gtctgcctgt 5280

tcatccgcgt ccagctcggt gagtttctcc agaagcgta atgtctggct tctgataaag 5340

cgggccatgt taagggcggt tttttcctgt ttggtcactg atgcctccgt gtaaggggga 5400

tttctgttca tgggggtaat gataccgatg aaacgagaga ggatgctcac gatacggggtt 5460

actgatgatg aacatgcccg gttactggaa cgttgtgagg gtaaacaact ggcggtatgg 5520

atgcggcggg accagagaaa aatcactcag ggtcaatgcc agcgcttcgt taatacagat 5580

gtaggtgttc cacagggtag ccagcagcat cctgcgatgc agatccggaa cataatggtg 5640

cagggcgctg acttccgcgt ttccagactt tacgaaacac ggaaaccgaa gaccattcat 5700

gttggtgctc aggtcgcaga cgttttgcag cagcagtcgc ttcacgttcg ctgcggtatc 5760

ggtgattcat tctgctaacc agtaaggcaa ccccgccagc ctagccgggt cctcaacgac 5820

aggagcacga tcatgcgac cctgggccag gacccaacgc tgcccagat gcgccgcgtg 5880

cggctgctgg agatggcgga cgcgatggat atgttctgcc aagggttggg ttgcgcattc 5940

acagttctcc gcaagaattg attggctcca attcttgag tggatgaatcc gttagcgagg 6000

tgccgccggc ttccattcag gtcgaggtgg cccggctcca tgcaccgga cgcaacgcgg 6060

ggaggcagac aaggtatagg gcggcgcta caatccatgc caaccgttc catgtgctcg 6120

ccgaggcggc ataaatcgcc gtgacgatca gcggtccagt gatcgaagtt aggctggtaa 6180

gagccgcgag cgatccttga agctgtccct gatggtcgtc atctacctgc ctggacagca 6240

tggcctgcaa cgcgggcatc ccgatgccgc cggaagcgag aagaatcata atggggaagg 6300

ccatccagcc tcgcgtcgcg aacgccagca agacgtagcc cagcgcgtcg gccgccatgc 6360

cggcgataat ggcttgcttc tcgccgaaac gtttggtggc gggaccagtg acgaaggctt 6420

gagcgagggc gtgcaagatt ccgaataccg caagcgacag gccgatcatc gtcgcgctcc 6480

agcgaaagcg gtcctcgccg aaaatgaccc agagcgtgc cggcacctgt cctacgagtt 6540

gcatgataaa gaagacagtc ataagtgcgg cgacgatagt catgccccgc gcccacccga 6600

aggagctgac tgggttgaag gctctcaagg gcatcggtcg agatcccggg gcctaattgag 6660

tgagctaact tacattaatt gcgttgcgct cactgccccgc tttccagtcg ggaaacctgt 6720

cgtgccagct gcattaatga atcggccaac gcgcggggag aggcggtttg cgtattgggc 6780

gccagggtagg tttttctttt caccagttag acgggcaaca gctgattgcc cttcacccgc 6840

tggccctgag agagttgcag caagcgggtc acgctggttt gccccagcag gcgaaaatcc 6900

tgtttgatgg tggttaacgg cgggatataa catgagctgt cttcggtatc gtcgtatccc 6960

actaccgaga tatccgcacc aacgcgcagc ccggactcgg taatggcgcg cattgcgccc 7020

agcgccatct gatcgttggc aaccagcatc gcagtgggaa cgatgccctc attcagcatt 7080

tgcattggttt gttgaaaacc ggacatggca ctccagtgcg cttcccgttc cgctatcggc 7140

tgaatttgat tgcgagtgcg atatattatgc cagccagcca gacgcagacg cgccgagaca 7200

gaacttaatg ggcccgctaa cagcgcgatt tgctggtgac ccaatgcgac cagatgctcc 7260

acgcccagtc gcgtaccgtc ttcattggag aaaataatac tgttgatggg tgtctggcca 7320

gagacatcaa gaaataacgc cggaacatta gtgcaggcag cttccacagc aatggcatcc 7380

tggtcatcca gcggatagtt aatgatcagc cactgacgc gttgcgcgag aagattgtgc 7440

accgccgctt tacaggcttc gacgcgcgtt cgttctacca tcgacaccac cacgctggca 7500

cccagttgat cggcgcgaga ttaatacgcc gcgacaattt gcgacggcgc gtgcagggcc 7560

agactggagg tggcaacgcc aatcagcaac gactgtttgc ccgccagttg ttgtgccacg 7620

cggttgggaa tgtaattcag ctccgccatc gccgcttcca ctttttcccg cgttttcgca 7680

gaaacgtggc tggcctggtt caccacgcgg gaaacggtct gataagagac accggcatac 7740

tctgcgacat cgtataacgt tactggtttc acattcacca ccctgaattg actctcttcc 7800

gggcgctatc atgccatacc gcgaaagggt ttgcgccatt cgatggtgtc cgggatctcg 7860

acgctctccc ttatgcgact cctgcattag gaagcagccc agtagtaggt tgaggccggt 7920

gagcaccgcc gccgcaagga atggtgcatg caaggagatg gcgccaaca gtcccccgcc 7980

cacggggcct gccaccatac ccacgccgaa acaagcgctc atgagcccga agtggcgagc 8040

ccgatcttcc ccatcggtga tgtcggcgat ataggcgcca gcaaccgcac ctgtggcgcc 8100

ggtgatgccg gccacgatgc gtccggcgta gaggatcgag atctcgatcc cgcgaaa 8157

<210> 4

<211> 7659

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Genetically
engineered plasmid containing helicase domain of
HCV NS3

<400> 4

ttaatacgac tcactatagg ggaattgtga gcggataaca attcccctct agaaataatt 60

ttgtttaact ttaagaagga gatataccat ggtggacttt atccctgtgg agaacctgga 120

gaccaccatg cgttccccgg tgttcacgga caactcctct ccaccagctg tccccagag 180

cttccaggtg gccacactgc atgctccac cggcagtggg aagagcacca aggtcccggc 240

tgcgtacgca gcccagggct acaagggtgtt ggtgctcaac ccctctgttg ctgcaacgct 300

gggcttttgt gcttacatgt ccaaggccca tggggtcgat cctaatatcc gcaccggtgt 360

gcgtacaatt accactggca gcccacac gtactccacc tacggcaagt tccttgccga 420

cggcgggtgc tcaggtggcg cttatgatat catcatttgt gacgagtgcc actccacgga 480

tgccacatcc atcttgggca tcggcactgt ccttgaccaa gcagagactg cgggggag 540

attggttggtg ctgcccactg ctacccctcc gggctccgtc acggtaccgc atcctaacaat 600

cgaggagggtt gctctgtcca ccaccggaga gatccctttc tacggcaagg ctatccccct 660

cgagggtgatc aaggggcgcc gtcattctcat cttctgtcac tcaaagaaga agtgcgacga 720

gctcgccgcg aagctggtcg cattgggcat caatgccgtg gcctactacc gcggacttga 780

cgtgtctgtc atcccgacca gcggcgatgt tgcgtcgtg gcgaccgatg ctctcatgac 840

tggctttacc ggcgacttcg actctgtgat agactgcaac acgtgtgtca ctacagacagt 900

cgatttcagc cttgacccta cctttaccat tgagacaacc acgctcccc aggatgctgt 960

ctccaggact cagcgccgtg gtcgtaccgg ccgtgggaag ccaggcatct acagatttgt 1020

ggcaccgggg gagcgccct ccggcatgtt cgactcgtcc gtcctctgtg agtgctatga 1080

cgcgggctgt gcttggtatg agctcacgcc ggcggagact acagtctgtc tgcgcgcgta 1140

catgaacacc ccggggcttc ccgtgtgcca ggaccatctt gaattttggg agggcgtctt 1200

tacgggcctc acccatatcg atgcccactt tctgtcccag acaaagcaga gtggggagaa 1260

ctttccttac ctggtagcgt accaagccac cgtgtgcgct cgtgcgcaag cccctccgcc 1320

atcgtgggac cagatgtgga agtgtttgat ccgccttaaa cccaccctcc atgggccaac 1380

accgctcctg taccgtctgg gcgctgttca gaatgaagtc accctgacgc acccaatcac 1440

caaatacatc atgacatgca tgtcggccga cctggaggtc gtcacgggat ctggctcgca 1500

tcatcatcat catcactaat agaattcgga tccggctgct aacaaagccc gaaaggaagc 1560

tgagttggct gctgccaccg ctgagcaata actagcataa ccccttgggg cctctaaacg 1620

ggtcttgagg ggttttttgc tgaaaggagg aactatatcc ggatatcccg caagaggccc 1680

ggcagtaccg gcataaccaa gcctatgcct acagcatcca ggtgacggt gccgaggatg 1740

acgatgagcg cattgttaga tttcatacac ggtgcctgac tgcgttagca atttaactgt 1800

gataaactac cgcattaaag cttatcgata ccgtcgacct cgaggggggg cccggtaccc 1860

aattcgccct atagtgagtc gtattacgog cgctcactgg ccgtcgtttt acaacgtcgt 1920

gactgggaaa accctggcgt tacccaactt aatcgccctg cagcacatcc ccctttcgcc 1980

agctggcgta atagcgaaga ggcccgaccc gatcgccctt cccaacagtt gcgcagcctg 2040

aatggcgaat gggacgcgcc ctgtagcggc gcattaagcg cggcgggtgt ggtggttacg 2100

cgcagcgtga ccgctacact tgccagcgcc ctacgccccg ctcttttcgc tttcttcctt 2160

tcctttctcg ccacgttcgc cggctttccc cgtcaagctc taaatcgggg gctcccttta 2220

gggttccgat ttagtgcttt acggcacctc gacccccaaa aacttgatta gggatgatgt 2280

tcacgtagtg ggccatcgcc ctgatagacg gtttttcgcc ctttgacgtt ggagtccacg 2340

ttctttaata gtggactctt gttccaaact ggaacaacac tcaaccctat ctcggtctat 2400

tccttttgatt tataagggat tttgcgatt tcggcctatt ggtaaaaaa tgagctgatt 2460

taacaaaaat ttaacgcgaa ttttaacaaa atattaacgc ttacaattta ggtggcactt 2520

ttcggggaaa tgtgcgcgga acccctattt gtttattttt ctaaatacat tcaaatatgt 2580

atccgctcat gagacaataa ccctgataaa tgcttcaata atattgaaaa aggaagagta 2640

tgagtattca acatttccgt gtcgccctta ttcccttttt tgcggcattt tgccttcctg 2700

tttttgctca ccagaaaacg ctggtgaaag taaaagatgc tgaagatcag ttgggtgcac 2760

gagtgggtta catcgaactg gatctcaaca gcggttaagat ccttgagagt ttccgccccg 2820

aagaacgttt tccaatgatg agcactttta aagttctgct atgtggcgcg gtattatccc 2880

gtattgacgc cgggcaagag caactcggtc gccgcataca ctattctcag aatgacttgg 2940

ttgagtactc accagtcaca gaaaagcatc ttacggatgg catgacagta agagaattat 3000

gcagtgctgc cataaccatg agtgataaca ctgcggccaa cttacttctg acaacgatcg 3060

gaggaccgaa ggagctaacc gcttttttgc acaacatggg ggatcatgta actcgccttg 3120

atcgttggga accggagctg aatgaagcca taccaaacga cgagcgtgac accacgatgc 3180

ctgcagcaat ggcaacaacg ttgcgcaaac tattaactgg cgaactactt actctagctt 3240

cccggcaaca attaatagac tggatggagg cggataaagt tgcaggacca cttctgcgct 3300

eggcccttcc ggctggctgg tttattgctg ataaatctgg agccggtgag cgtgggtctc 3360

gcggtatcat tgcagcactg gggccagatg gtaagccctc ccgtatcgta gttatctaca 3420

cgacggggag tcaggcaact atggatgaac gaaatagaca gatcgctgag ataggtgcct 3480

cactgattaa gcattggtaa ctgtcagacc aagtttactc atatatactt tagattgatt 3540

taaaacttca tttttaattt aaaaggatct aggtgaagat cttttttgat aatctcatga 3600

ccaaaatccc ttaacgtgag ttttcgttcc actgagcgtc agaccccgta gaaaagatca 3660

aaggatcttc ttgagatcct ttttttctgc gcgtaatctg ctgcttgcaa acaaaaaaac 3720

caccgctacc agcggtggtt tgtttgcgg atcaagagct accaactctt tttccgaagg 3780

taactggctt cagcagagcg cagataccaa atactgtcct tctagtgtag ccgtagttag 3840

gccaccactt caagaactct gtagcacgcg ctacatacct cgctctgcta atcctgttac 3900

cagtggctgc tgccagtggc gataagtcgt gtcttaccgg gttggactca agacgatagt 3960

taccggataa ggcgcagcgg tcgggctgaa cgggggggttc gtgcacacag cccagcttgg 4020

agcgaacgac ctacaccgaa ctgagatacc tacagcgtga gctatgagaa agcgccacgc 4080

ttcccgaagg gagaaaggcg gacaggtatc cggtaagcgg cagggtcggg acaggagagc 4140

gcacgagggg gcttccaggg ggaaacgcct ggtatcttta tagtcctgtc gggtttcgcc 4200

acctctgact tgagcgtcga tttttgtgat gtcgctcagg ggggcggagc ctatggaaaa 4260

acgccagcaa cgcggccttt ttacggttcc tggccttttg ctggcctttt gtcacatgt 4320

tctttcctgc gttatccctt gattctgtgg ataaccgtat taccgccttt gaggtagctg 4380

ataccgctcg ccgcagccga acgaccgagc gcagcgagtc agtgagcgag gaagcggaag 4440

agcgctgat gcggtatattt ctccttacgc atctgtgcgg tatttcacac cgcataatg 4500

gtgcactctc agtacaatct gctctgatgc cgcatagtta agccagtata cactccgcta 4560

tcgctacgtg actgggtcat ggctgcgccc cgacacccgc caacacccgc tgacgcgccc 4620

tgacgggctt gtctgtctcc ggcatacgtt tacagacaag ctgtgaccgt ctccgggagc 4680

tgcattgtgc agaggttttc accgtcatca ccgaaacgcg cgaggcagct gcggtaaagc 4740

tcattcagct ggctcgtgaag cgattcacag atgtctgcct gttcatccgc gtccagctcg 4800

ttgagtttct ccagaagcgt taatgtctgg cttctgataa agcggggccat gttaagggcg 4860

gttttttctt gtttggtcac tgatgcctcc gtgtaagggg gatttctgtt catgggggta 4920

atgataccga tgaaacgaga gaggatgctc acgatacggg ttactgatga tgaacatgcc 4980

cggttactgg aacgttgtga gggtaaacia ctggcggtat ggatgcggcg ggaccagaga 5040

aaaatcactc aggggtcaatg ccagcgcttc gttaatacag atgtaggtgt tccacagggt 5100

agccagcagc atcctgcat gcagatccgg aacataatgg tgcagggcgc tgacttccgc 5160

gtttccagac ttacgaaac acggaaaccg aagaccattc atgttggtgc tcaggtcgca 5220

gacgttttgc agcagcagtc gcttcacgtt cgctcgcgta tcggtgattc attctgctaa 5280

ccagtaaggc aaccccgcca gcctagccgg gtcctcaacg acaggagcac gatcatgcgc 5340

accgtggcc aggacccaac gctgcccag atgcgccgcg tgcggctgct ggagatggcg 5400

gacgcgatgg atatgttctg ccaagggttg gtttgcgcac tcacagttct ccgcaagaat 5460

tgattggctc caattcttgg agtggtgaat ccgtagcga ggtgccgccg gcttccattc 5520

aggtcgaggt ggcccggctc catgcaccgc gacgcaacgc ggggaggcag acaaggtata 5580

gggcggcgcc tacaatccat gccaaaccgt tccatgtgct cgccgaggcg gcataaatcg 5640

ccgtgacgat cagcgtcca gtgatcgaag ttaggctggt aagagccgcg agcgatcctt 5700

gaagctgtcc ctgatggtcg tcatctacct gcctggacag catggcctgc aacgcgggca 5760

tcccgatgcc gccggaagcg agaagaatca taatggggaa ggccatccag cctcgcgctcg 5820

cgaacgccag caagacgtag cccagcgcggt cgcccgccat gccggcgata atggcctgct 5880

tctcgccgaa acgtttgggtg gcgggaccag tgacgaaggc ttgagcgagg gcgtgcaaga 5940

ttccgaatac cgcaagcgac aggccgatca tcgtcgcgct ccagcgaaaag cggtcctcgc 6000

cgaaaatgac ccagagcgct gccggcacct gtcctacgag ttgcatgata aagaagacag 6060

tcataagtgc ggcgacgata gtcatgcccc gcgcccaccg gaaggagctg actgggttga 6120

aggctctcaa gggcatcggt cgagatcccc gtgcctaata agtgagctaa cttacattaa 6180

ttgcgttgcg ctactgccc gctttccagt cgggaaacct gtcgtgccag ctgcattaat 6240

gaatcggcca acgcgcgggg agaggcggtt tgcgtattgg gcgccagggt gggtttttctt 6300

ttcaccagtg agacgggcaa cagctgattg cccttcaccg cctggcctg agagagttgc 6360

agcaagcggt ccacgctggt ttgccccagc aggcgaaaaat cctgtttgat ggtgggttaac 6420

ggcgggatat aacatgagct gtcttcggta tcgtcgtatc ccactaccga gatatccgca 6480

ccaacgcgca gcccggaactc ggtaatggcg cgcattgcgc ccagcgccat ctgatcggtg 6540

gcaaccagca tcgcagtggg aacgatgcc tcattcagca ttgcatggt ttgttgaaaa 6600

ccggacatgg cactccagtc gccttcccgt tccgctatcg gctgaatttg attgcgagtg 6660

agatatattat gccagccagc cagacgcaga cgcgcgcgaga cagaacttaa tgggcccgct 6720

aacagcgcgga ttgctgggtg acccaatgcg accagatgct ccacgcccag tcgcgtaccg 6780

tcttcatggg agaaaataat actgttgatg ggtgtctggt cagagacatc aagaaataac 6840

gccggaacat tagtgcaggc agcttcaca gcaatggcat cctgggtcacc cagcggatag 6900

ttaatgatca gccactgac gcgttgcgcg agaagattgt gcaccgccgc ttacaggt 6960

tcgacgccgc ttcgttctac catcgacacc accacgttg caccagttg atcggcgcga 7020

gatttaatcg ccgcgacaat ttgcgacggc gcgtgcaggc ccagactgga ggtggcaacg 7080

ccaatcagca acgactgttt gcccgccagt tgttgtgcca cgcggttggg aatgtaattc 7140

agctccgcca tcgccgcttc cactttttcc cgcgttttcg cagaaacgtg gctggcctgg 7200

ttcaccacgc gggaaacggt ctgataagag acaccggcat actctgcgac atcgtataac 7260

gttactgggtt tcacattcac caccctgaat tgactctctt ccgggcgcta tcatgccata 7320

ccgcgaaagg ttttgcgcca ttcgatgggtg tccgggatct cgacgctctc ccttatgcga 7380

ctctgcatt aggaagcagc ccagtagtag gttgaggccg ttgagcaccg ccgccgcaag 7440

gaatggtgca tgcaaggaga tggcgcccaa cagtcccccg gccacggggc ctgccaccat 7500

accacgccc aaacaagcgc tcatgagccc gaagtggcga gccgatctt ccccatcggt 7560

gatgtcggcg atataggcgc cagcaaccgc acctgtggcg ccggtgatgc cggccacgat 7620

gcgtccggcg tagaggatcg agatctcgat cccgcgaaa

7659